# Introduction to Deep Learning (I2DL)

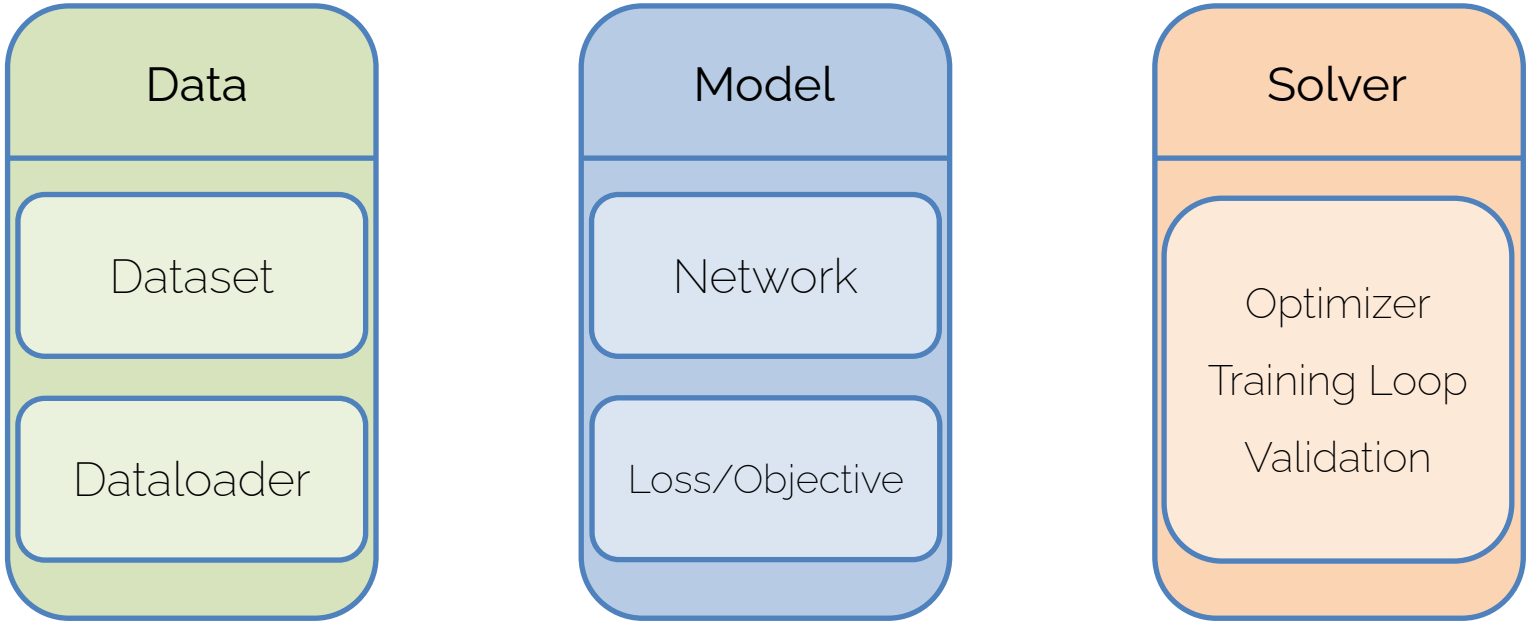
## Exercise 4: Simple Classifier

# Today's Outline

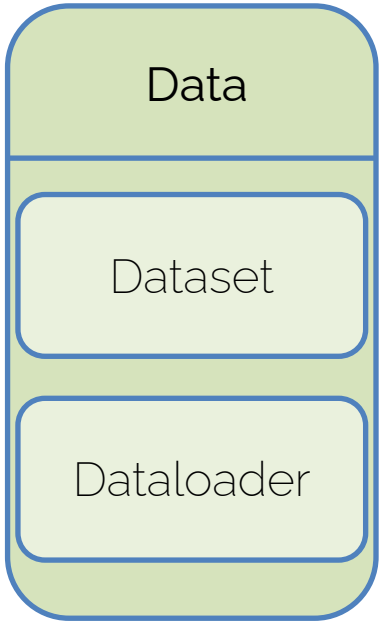
- The Pillars of Deep Learning
- Exercise 4: Simple Classifier  Binary Prediction
  - Housing Dataset
  - Training loop: Forward & Backward pass
- Backpropagation

# The Pillars of Deep Learning

# The Pillars of Deep Learning



# The Pillars of Deep Learning



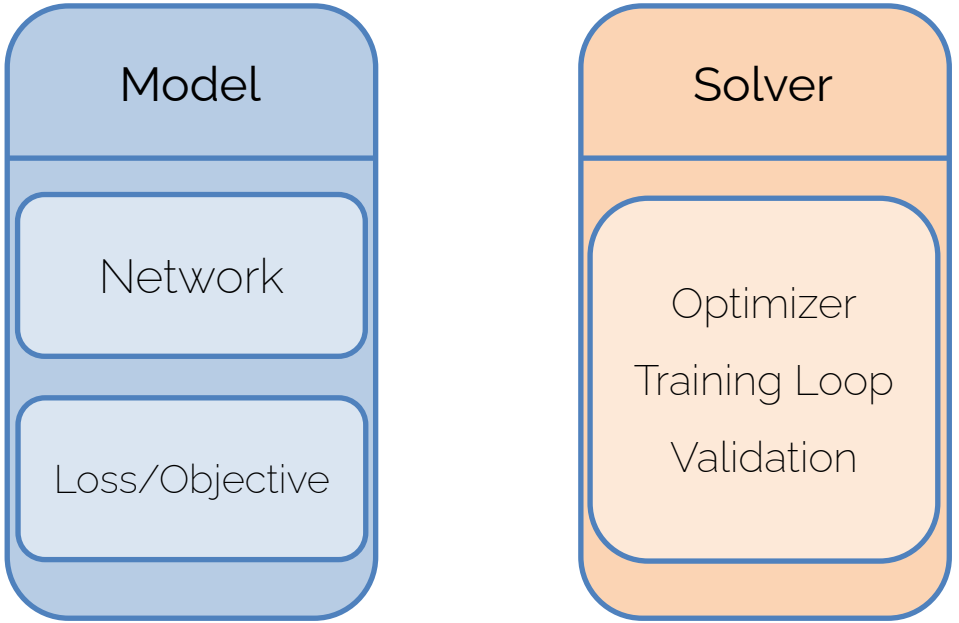

Exercise 3: Dataset and Dataloader

# The Pillars of Deep Learning

Exercise 4: Simple Classifier

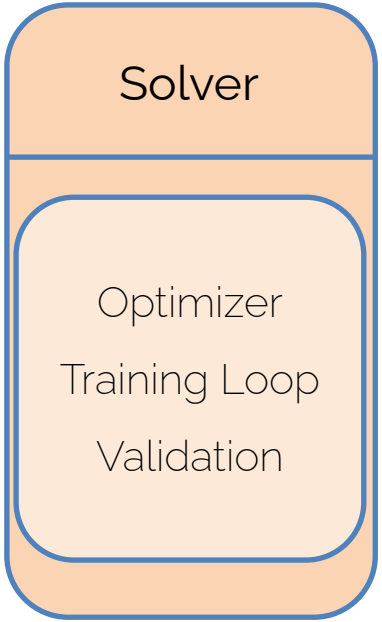
Exercise 5: Simple Network

Exercise 6: Hyperparameter Tuning



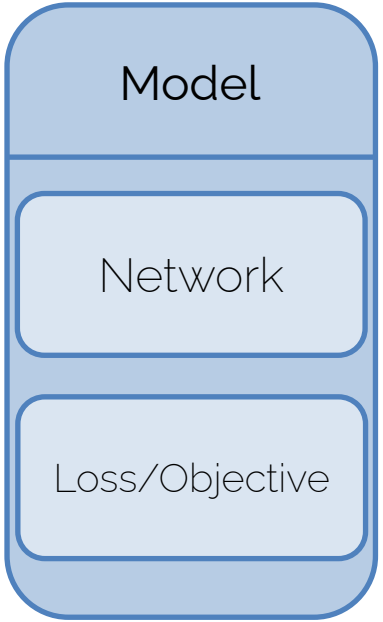
# Goal: Exercise 4

- Goal: Training process
- Skip: Model Pillar
- Simplified Model: Classifier which is a 1-Layer Neural Network



# Goals: Exercises 5++

- Ex 3 + 4: Dataloading and Trainings process
- Ex 5++: Expand the exercises to more interesting model architectures



# Exercise 4: Simple Classifier

# Housing Dataset

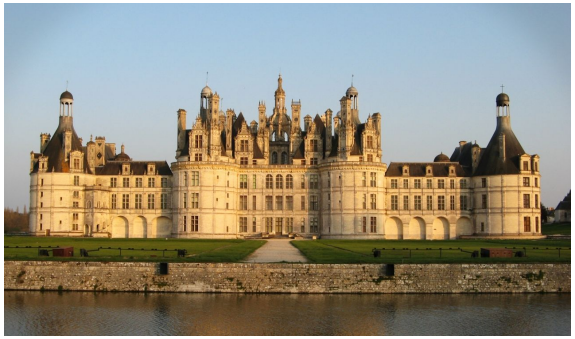
- **Housing Dataset**: Data of ~1400 houses including 81 features like Neighborhood, GrLivArea, YearBuilt, etc.

- **Simplified model**: 1 input feature to predict house price label ("expensive" vs "low-prized")

housing_train

| Id | Neighborhood | BldgType | HouseStyle | YearBuilt | YearRemodAdd | RoofStyle | CentralAir | GrLivArea | FullBath | HalfBath | Fireplaces | PoolArea | Fence | SalePrice |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 1 | CollgCr | 1Fam | 2Story | 2003 | 2003 | Gable | Y | 1710 | 2 | 1 | 0 | 0 | NA | 208500 |
| 2 | Veenker | 1Fam | 1Story | 1976 | 1976 | Gable | Y | 1262 | 2 | 0 | 1 | 0 | NA | 181500 |
| 3 | CollgCr | 1Fam | 2Story | 2001 | 2002 | Gable | Y | 1786 | 2 | 1 | 1 | 0 | NA | 223500 |
| 4 | Crawfor | 1Fam | 2Story | 1915 | 1970 | Gable | Y | 1717 | 1 | 0 | 1 | 0 | NA | 140000 |
| 5 | NoRidge | 1Fam | 2Story | 2000 | 2000 | Gable | Y | 2198 | 2 | 1 | 1 | 0 | NA | 250000 |
| 6 | Mitchel | 1Fam | 1.5Fin | 1993 | 1995 | Gable | Y | 1362 | 1 | 1 | 0 | 0 | MnPrv | 143000 |
| 7 | Somerst | 1Fam | 1Story | 2004 | 2005 | Gable | Y | 1694 | 2 | 0 | 1 | 0 | NA | 307000 |
| 8 | NWAmes | 1Fam | 2Story | 1973 | 1973 | Gable | Y | 2090 | 2 | 1 | 2 | 0 | NA | 200000 |

# Exercise 4 - Classifying House Prices



ML Model $M$
$M(\mathbf{x}) = \mathbf{y}$

Expensive $\boldsymbol{y = 1}$

ML Model $M$
$M(\mathbf{x}) = \mathbf{y}$

Low-priced $\boldsymbol{y = 0}$

# 3rd Pillar of Deep Learning

Data

Training Data

Validation Data

Model

Model

Loss Function

Solver

Training Loop → Forward Pass, Backward Pass

Optimizer → Gradient Descent

Validation

Exercise 03

Very simple model

# Backpropagation

# Backpropagation: Overview

Forward pass

$\mathbf{x}$

Input

$M_\theta(x) = \hat{y}$

$d\hat{y}/d\theta$

$\hat{y}$

Prediction

$L(y, \hat{y})$

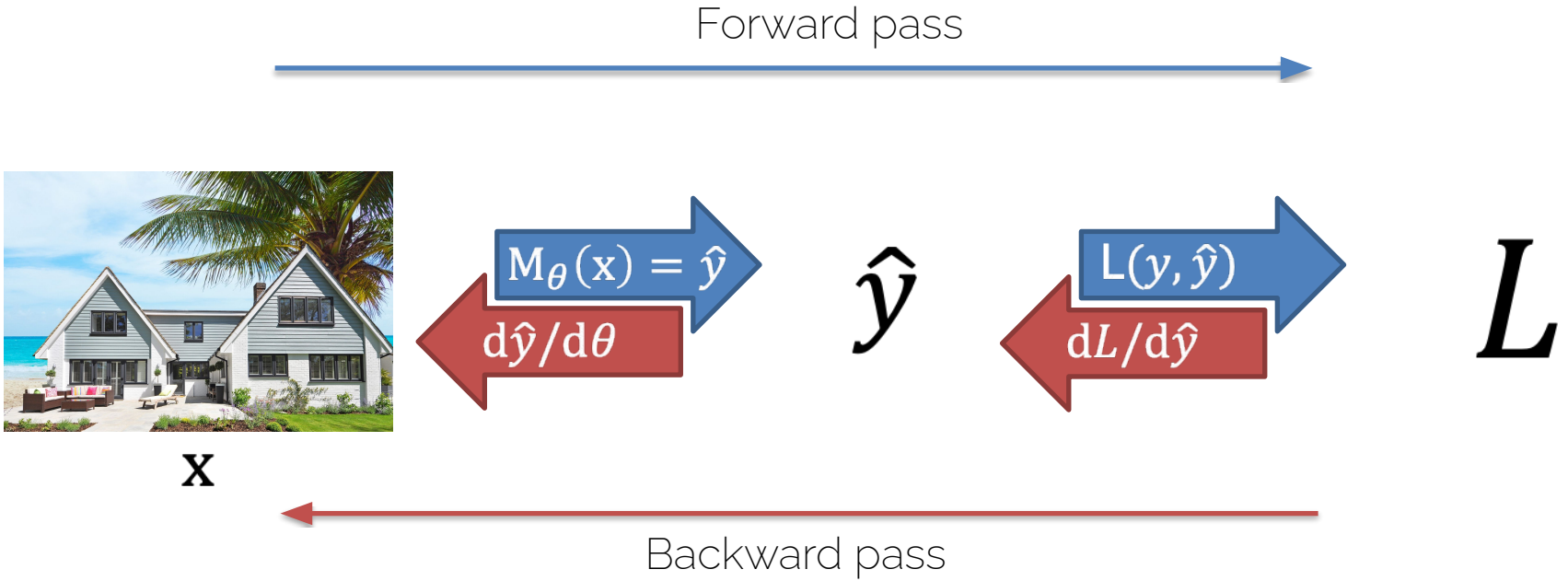
$dL/d\hat{y}$

$L$

Loss

Backward pass

# Backpropagation: Loss Function




Binary Cross Entropy Loss:

$$L(y, \hat{y}) = y \cdot log(\hat{y}) + (1 - y) \cdot log(1 - \hat{y})$$

# Backpropagation: Update Step

Forward pass

$\mathbf{x}$

$M_\theta(x) = \hat{y}$

$d\hat{y}/d\theta$

$\hat{y}$

$L(y, \hat{y})$

$dL/d\hat{y}$

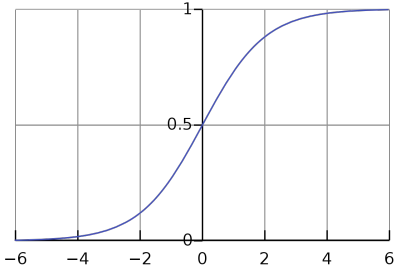
$L$

Backward pass

Optimization with gradient descent: $\theta_{t+1} = \theta_t - \lambda \cdot \nabla_\theta L$

# Backpropagation: Summary

- Input: $X \in \mathbb{R}^{N \times D+1}$ representing our data with N samples and D+1 feature dimensions
- Output: Binary labels given by $y \in \mathbb{R}^{N \times 1}$
- Model: Classifier of the form $y = \sigma(X \cdot w)$
- Sigmoid function: $\sigma : \mathbb{R} \rightarrow [0, 1]$ with $\sigma(t) = \frac{1}{1+e^{-t}}$



$\sigma$ Function

- Weights of the Classifier: $w = (w_1, w_2, \ldots, w_{D+1}) \top \in \mathbb{R}^{D+1}$

# Backpropagation: Example

# Backpropagation

Forward pass

$\mathbf{x}$

$M_\theta(x) = \hat{y}$

$d\hat{y}/d\theta$

$\hat{y}$

$L(y, \hat{y})$

$dL/d\hat{y}$

$L$

Backward pass

# Forward Pass

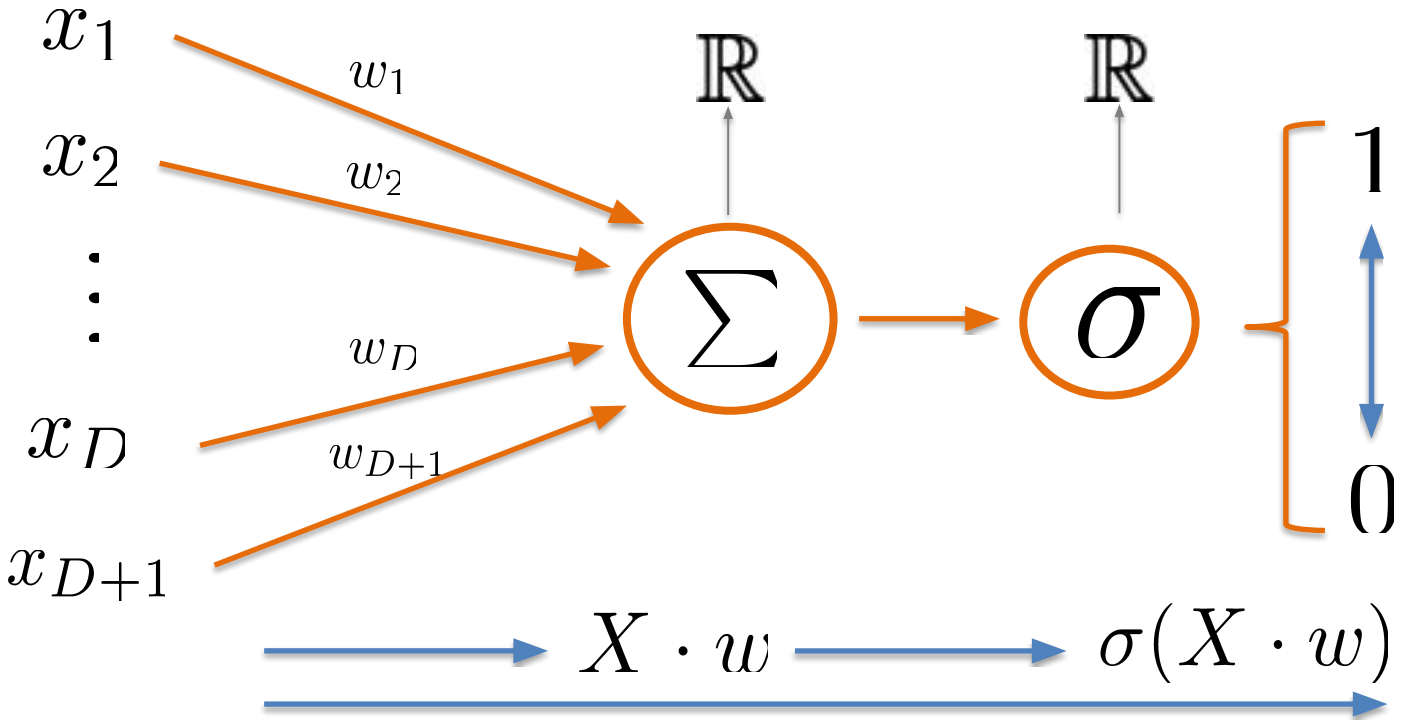
(Single sample)

Sample

$x = (x_1, x_2, \ldots, x_{D+1})$

Classifier Model

$y = \sigma(X \cdot w)$



# Input Data $X$

(Single sample -> N samples)

$$X \in \mathbb{R}^{N \times D+1}$$

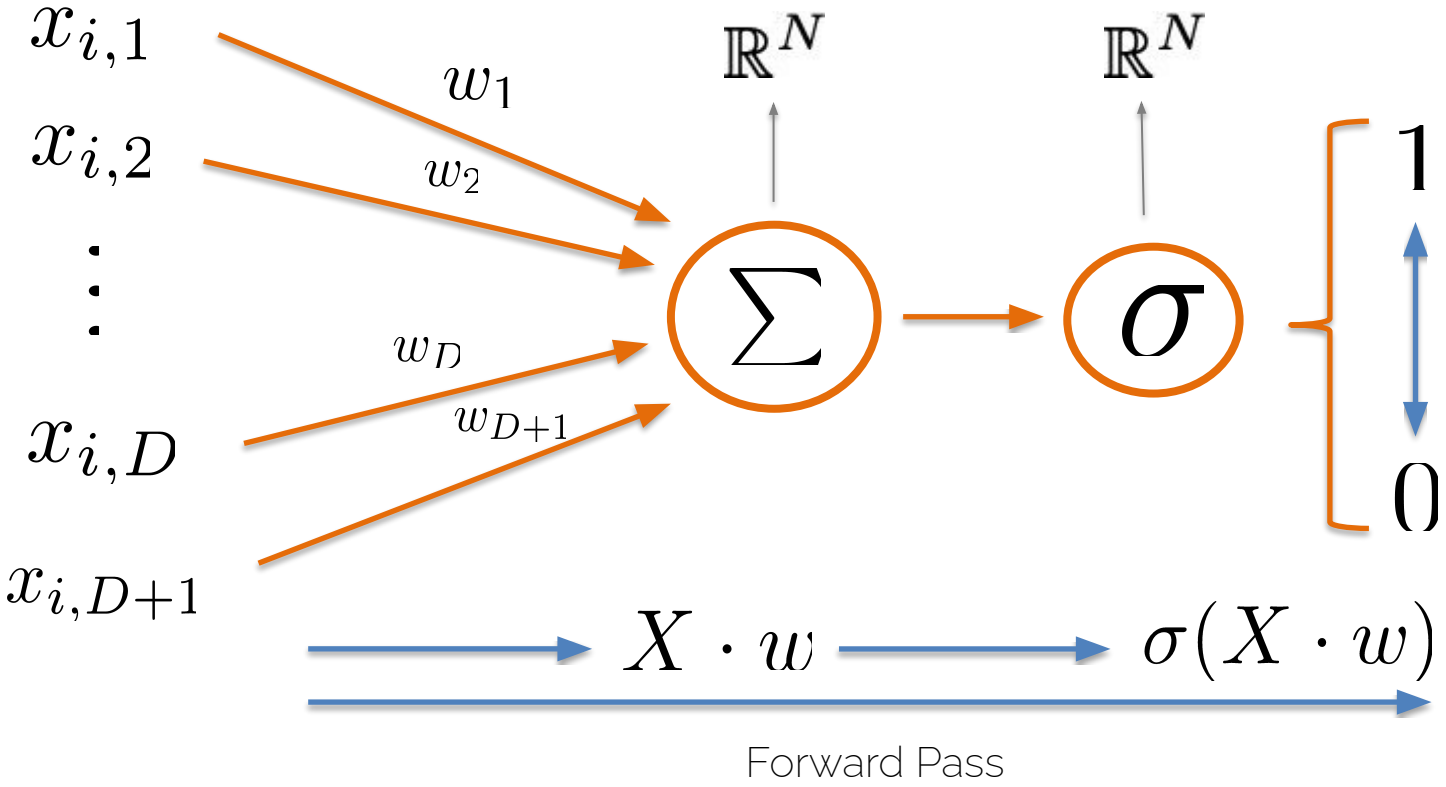
$$X = \begin{pmatrix} x_{1,1} & x_{1,2} & \dots & x_{1,D+1} \\ x_{2,1} & x_{2,2} & \dots & x_{2,D+1} \\ \vdots & \vdots & \ddots & \vdots \\ x_{N,1} & x_{N,2} & \dots & x_{N,D+1} \end{pmatrix}$$

# Forward Pass

(N samples)

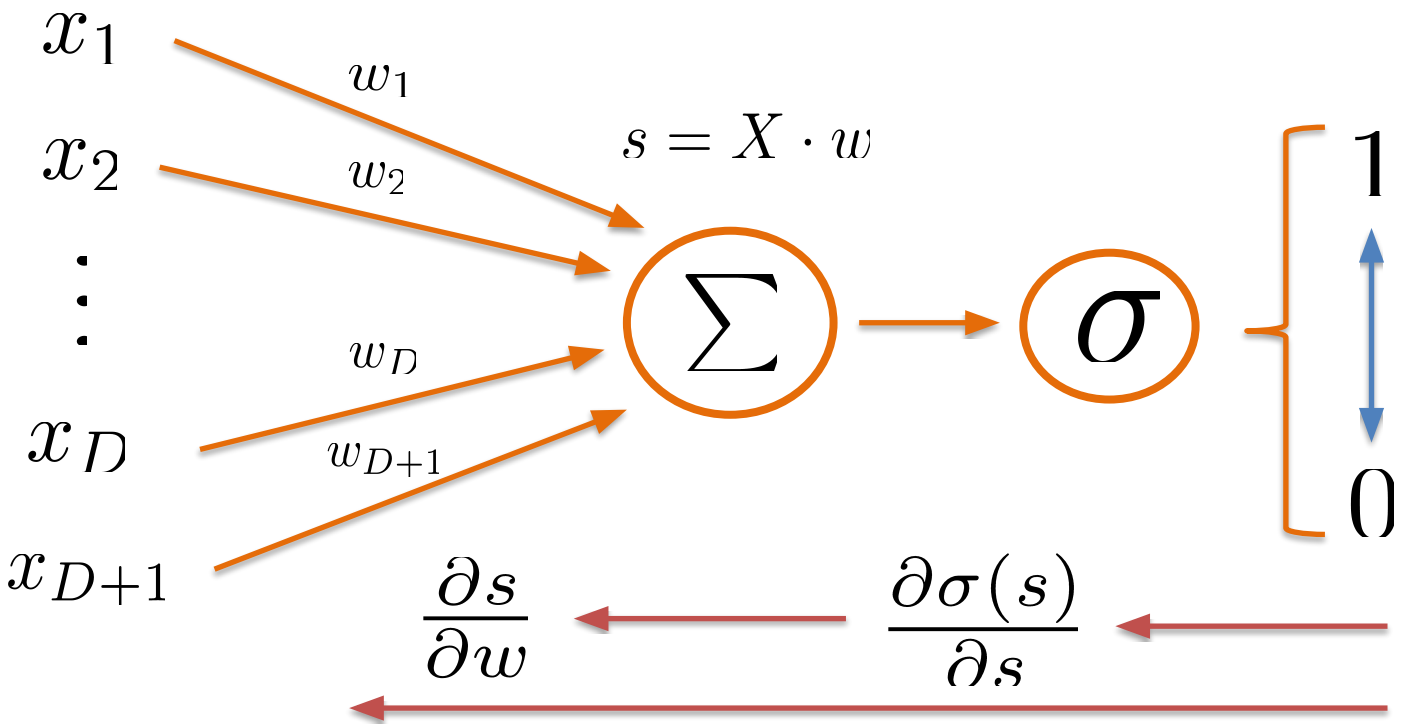
Sample

$x_i = (x_{i1}, x_{i2}, \ldots, x_{i,D+1})$



# Backward Pass

Sample
$x = (x_1, x_2, \ldots, x_{D+1})$




Backward Pass

# Backward Pass

- **Backward Pass**: Derivative of function with respect to weights

$$w = (w_1, w_2, \ldots, w_{D+1})$$ of our Classifier

- **Attention**: Make sure you understand the dimensions here
- **Step 1**: Forward + Backward Pass for one sample
- **Step 2**: Forward + Backward Pass for N samples

# Overview Exercise 4

- Two Notebooks
  - Optional: Preprocessing
  - Logistic regression model

Fixed Deadline:
May 28, 2025 23:59

- Submission
  - Several implementation tasks in the notebook
  - Submission file creation in Notebook

# See you next week